Выпуск 45(62) О П Т И М И З А Ц И Я 1989 г.

Численные методы оптимизации

УДК 519.8

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ НЕТОЧНЫМ МЕТОДОМ НЬЮТОНА ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНОЙ ЗАДАЧИ ДОПОЛНИТЕЛЬНОСТИ

Г.Ю.Грязина

Рассмотрим нелинейную задачу дополнительности: найти вектор $x \in R^n$ такой, что

$$x \geq 0,\ f(x) \geq 0,\ x^T f(x) = 0, \tag{1}$$

где $f: R^n \to R^n$ – отображение класса $C^1(R^n)$. В работах [1] и [2] изучены условия, гарантирующие существование и единственность решения задачи (1). Так, например, если отображение f сильно монотонно, то существует единственное решение x^* данной задачи.

Для практических целей часто требуется найти приближенное решение задачи (1). Для его нахождения применяются различные итеративные методы, многие из которых рассмотрены в работе [3]. Одним из них является метод Ньютона, решающий задачу (1) путем построения последовательности точек $\{x^k\}, k = 0, 1, \ldots$, где x^0 – начальное приближение решения, а x^{k+1} получается из x^k как решение линейной задачи дополнительности:

$$x \geq 0,\ W_k(x) = f(x^k) + \nabla f(x^k)(x - x^k) \geq 0,\ x^T W_k(x) = 0. \tag{2}$$

Вопрос о сходимости последовательности приближений $\{x^k\}$ к решению x^* задачи (1) также исследуется в работе [3]. В частности, там доказывается результат о локальной сходимости метода (2).

В настоящей работе задача дополнительности (1) решается неточным методом Ньютона, который позволяет в качестве x^{k+1}

использовать приближенное решение задачи (2). Для нахождения такого решения необходим эффективно вычислимый критерий близости вектора x^{k+1} к точному решению задачи (2). С целью получения такого критерия в работе [4] рассматриваются функции $h_k(x) = min(x; w_k(x))$ и $h(x) = min(x; f(x))$, а в [5] показывается, что $\|h_k(x)\|(\|h(x)\|)$ может служить мерой близости вектора x к точному решению задачи (2) (задачи (I)). Именно, в работе [4] неточный метод Ньютона для задачи (I) сформулирован следующим образом. Если x^k - текущая итерация, то в качестве x^{k+1} выбирается любой элемент из R^n, удовлетворяющий условию

$$\|h_k(x^{k+1})\| \leq \delta_k, \tag{3}$$

где δ_k - заранее выбранный скаляр. При некоторых предположениях относительно отображения f и выборе скаляра δ_k в виде $\delta_k = \xi_k \|h(x^k)\|$ в [4] приводятся условия на величины ξ_k, $k = 0,1,\dots$, обеспечивающие локальную сходимость последовательности $\{x^k\}$ к x^* с квадратичной скоростью. В [5] при более сильных предположениях относительно отображения f выделены условия на скаляры ξ_k, при выполнении которых имеет место глобальная сходимость метода.

В указанных подходах предполагается априорный выбор последовательности $\{\xi_k\}$, определяющей точность выполнения внутренних шагов. В настоящей работе рассматривается вопрос об оптимизации этой последовательности. Целью является нахождение последовательности $\{\xi_k\}$, обеспечивающей минимальную трудоемкость решения задачи (I).

В п.I работы ставится задача оптимизации для определения управляющих параметров $\delta_k > 0$, $k = 0,1,\dots$, участвующих в неравенстве (3), и показывается, что решение этой задачи существует. В п.2 обосновывается практическая реализация отыскания управляющих параметров δ_k, включающая пересчет величин, от которых δ_k зависит.

И, наконец, в п.3 приводятся численные результаты управления неточным методом Ньютона, которые для используемых тестовых функций показывают эффективность управления.

I. Постановка и исследование задачи оптимизации

Мы рассматриваем неточный метод Ньютона, который строит последовательность приближений $\{x^k\}$ следующим образом: имея текущую итерацию x^k, вектор x^{k+1} находится в соответствии с правилом

$$\left.\begin{aligned} &\|h_k(x^{k+1})\| \le \gamma_k = \xi_k \|h(x^k)\|,\ \xi_k > 0, \\ &h_k(x) = \min(x; f(x^k) + \nabla f(x^k)(x - x^k)). \end{aligned}\right\} \tag{4}$$

Понятно, что общая трудоемкость процесса нахождения x^* зависит от величины параметров γ_k : чем меньше γ_k, тем больше объем вычислений на k-м шаге. С другой стороны, при неоправданно больших значениях γ_k возрастает число шагов, необходимых для достижения заданной точности ε, что также может привести к возрастанию общей трудоемкости. Поэтому для достижения минимальной трудоемкости необходимо решать задачу оптимизации, находящую значения управляющих параметров γ_k. Чтобы поставить такую задачу, необходимо найти связь между $\eta_k = \|h(x^k)\|$ - погрешностью перед началом k-го шага, γ_k - точностью выполнения k-го шага и $\eta_{k+1} = \|h(x^{k+1})\|$ - погрешностью после окончания этого шага. Получим эту связь в виде оценки $\eta_{k+1} \le \Psi(\eta_k, \gamma_k)$.

В работе [5] для функции

$$f(x) = Ax + \varphi(x), \tag{5}$$

где A - положительно определенная матрица с параметром μ, отображение $\varphi: R^n \to R^n$ из класса $C^2(R^n)$ такое, что $\|\nabla\varphi(x)\| \le q\mu$ для некоторого $0 < q < \frac{1}{3}$ и всех x из R^n, получена следующая оценка:

$$\eta_{k+1} \le L_k \eta_k^2 + M_k \gamma_k, \tag{6}$$

где

$$L_k = \frac{(1 + \|A\| + q\mu)^3 \tau^k}{\mu^3 (1-q)^3},$$

$$M_k = \frac{(1 + \|A\| + \|\nabla\varphi(x^k)\|)(1 + \|A\| + q\mu)}{\mu - \|\nabla\varphi(x^k)\|},$$

$r^{\kappa} = \sup\limits_{x \in V^{\kappa}} \|\varphi''(x)\|$, V^{κ} – шар с центром в x^* радиусом $\|x^{\kappa} - x^*\|$. Оценка (6), очевидно, имеет смысл лишь в некоторой окрестности решения x^*, именно, когда $\rho_{\kappa} < 1/L_{\kappa}$.

Докажем, что имеет место глобальное убывание функции ошибки ρ_{κ}. Для доказательства понадобятся следующие соотношения (см., например, [5]) для сильно монотонной функции $f(x)$:

$$a\|x-y\| \leq \|h(x)-h(y)\| \leq b\|x-y\|, \tag{7}$$

где $a = \frac{\nu}{L_V + 1}$, $b = L_V + 1$, ν – константа сильной монотонности для функции $f(x)$, $L_V = \max\limits_{x \in V} \|\nabla f(x)\|$ для некоторого выпуклого ограниченного множества $V \subset R^n$ такого, что $x, y \in V$.

УТВЕРЖДЕНИЕ. Пусть $f(x)$ – функция вида (5), где

$$0 < q < \frac{\sqrt{(4\|A\|+5)^2 + 16\mu} - (4\|A\|+5)}{8\mu} .$$

Тогда существует $\rho < 1$, при котором для последовательности приближений $\{x^{\kappa}\}$, построенной в соответствии с правилами (4), справедлива следующая оценка:

$$\rho_{\kappa+1} \leq \rho\, \rho_{\kappa} + M_{\kappa} \gamma_{\kappa} . \tag{8}$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим функции $\chi(x) = x - h(x)$ и $\tau(x) = f(x) - h(x)$. Очевидно, $\chi(x) \geq 0$, $\tau(x) \geq 0$, $(\chi(x), \tau(x)) = 0$ для любого $x \in R^n$. Для $y^{\kappa+1}$ – точного решения линейной задачи дополнительности (2) имеем

$$y^{\kappa+1} \geq 0, \quad w_{\kappa}(y^{\kappa+1}) = Ay^{\kappa+1} + \varphi(x^{\kappa}) + $$
$$+ \nabla\varphi(x^{\kappa})(y^{\kappa+1} - x^{\kappa}) \geq 0, \quad (y^{\kappa+1})^T w_{\kappa}(y^{\kappa+1}) = 0.$$

Тогда

$$0 \geq (y^{\kappa+1} - \chi(y^{\kappa+1}))^T \cdot (w_{\kappa}(y^{\kappa+1}) - \tau(y^{\kappa+1})) =$$
$$= h(y^{\kappa+1})^T [\varphi(x^{\kappa}) - \varphi(y^{\kappa+1}) + \nabla\varphi(x^{\kappa})(y^{\kappa+1} - x^{\kappa}) + h(y^{\kappa+1})].$$

Отсюда $\|h(y^{\kappa+1})\|^2 \leq \|h(y^{\kappa+1})\| \cdot \|\varphi(y^{\kappa+1}) - \varphi(x^{\kappa}) - \nabla\varphi(x^{\kappa})(y^{\kappa+1} - x^{\kappa})\|$.

Так как $\|\nabla\varphi(x)\| \leq q\mu$, то $\|\varphi(y^{\kappa+1}) - \varphi(x^{\kappa})\| \leq q\mu \|y^{\kappa+1} - x^{\kappa}\|$, поэтому $\|h(y^{k+1})\| \leq 2q\mu \|y^{k+1} - x^k\|$. Далее, как показано в [5], функция $f(x)$ является сильно монотонной, поэтому можно воспользоваться соотношением (7), применяя которое к точкам x^k, y^{k+1}, получаем

$$\|h(y^{k+1})\| \leq 2q\mu \frac{1 + \|A\| + q\mu}{\mu(1-q)} \left(\|h(y^{k+1})\| + \|h(x^k)\|\right).$$

Откуда $\|h(y^{k+1})\| \leq \rho \|h(x^k)\|$, где $\rho = \dfrac{2q(1+\|A\|+q\mu)}{1-q(3+2\|A\|+2q\mu)}$.

Решив квадратное неравенство относительно q, можно показать, что для q, удовлетворяющего условию утверждения, $\rho < 1$. Далее, для x^{k+1}, построенного по правилам (4), имеем

$$\|h(x^{k+1})\| \leq \|h(y^{k+1})\| + \|h(x^{k+1}) - h(y^{k+1})\| \leq$$

$$\leq \rho \|h(x^k)\| + (1 + \|A\| + q\mu)\|x^{k+1} - y^{k+1}\| \leq$$

$$\leq \rho \|h(x^k)\| + \frac{(1+\|A\|+q\mu)(1+\|A\|+\|\nabla\varphi(x^k)\|)}{\mu - \|\nabla\varphi(x^k)\|} \|h_k(x^{k+1})\|,$$

что доказывает справедливость оценки (8).

ЗАМЕЧАНИЕ. В работе [5] для точных значений $\|\bar{h}(x)\| = \|x - x^*\|$ и $\|\bar{h}_k(x)\| = \|x - y^{k+1}\|$ доказывается аналогичное (8) соотношение при следующих ограничениях на q : $0 < q < \frac{1}{3}$. Покажем, что ограничение на q, требуемое в утверждении, ненамного жестче указанного.

Действительно, для $\mu << \|A\|^2$ (плохо обусловленных задач) имеем

$$0 < q < \frac{\sqrt{(4\|A\|+5)^2 + 16\mu} - (4\|A\|+5)}{8\mu} =$$

$$= \frac{4\|A\|+5}{8\mu}\left(\sqrt{1 + \frac{\mu}{\|A\|^2 + 5/2\|A\| + 25/16}} - 1\right) \approx$$

$$\approx \frac{4\|A\|+5}{8\mu} \cdot \frac{\mu}{2\|A\|^2 + 5\|A\| + 25/8} = \frac{1}{4\|A\|+5}.$$

Откуда для $\|A\| \approx 0$ получаем $0 < q \lesssim \frac{1}{5}$.

Из оценок (6) и (8) получаем, наконец, искомое соотношение

$$\eta_{k+1} \leq \psi(\eta_k, \delta_k) = \min\{\rho; L_k\eta_k\}\eta_k + M_k\delta_k, \qquad (9)$$

где $\rho < 1,\ L_k > 0,\ M_k > 0$. Далее предполагаем, что для заданной точности ε существует номер s , при котором $\eta_s \le \varepsilon$. Пусть m - наименьший из таких индексов. Тогда задача оптимизации для определения управляющих параметров $\delta_k > 0,\ k = 0,1,\ldots,$ ставится следующим образом:

$$\left.\begin{aligned}
&\min_{\{\delta_k\}_{k=0}^{m-1}} \left\{ \max_{\{\eta_k\}_{k=0}^{m-1}} \sum_{k=0}^{m-1} C(\eta_k,\delta_k),\ m \ge 1,\ \delta_k \ge 0,\ k=0,\ldots,m-1 \right\};\\
&\eta_{k+1} \le \min\{\rho;\ L_k\eta_k\}\eta_k + M_k\delta_k,\ k=0,\ldots,m-1;\\
&\eta_m \le \varepsilon.
\end{aligned}\right\} \qquad (10)$$

Заметим, что константы L_k и M_k, $k=0,1,\ldots$, можно оценить величинами L и M , не зависящими от шага k , именно,

$$L_k \le L = \frac{(1+\|A\|+q\mu)^3 r}{\mu^3(1-q)^3},\quad M_k \le M = \frac{(1+\|A\|+q\mu)^2}{\mu(1-q)},$$

где $r = \sup\limits_{x\in V^0} \|\varphi''(x)\|$, и поставить задачу оптимизации с постоянными константами, что упростило бы ее решение. Однако константы L и M при практической реализации, как правило, неизвестны, поэтому пришлось бы пользоваться их оценками, уточняющимися в процессе построения последовательности точек $\{x^k\}$, что привело бы к неизбежному изменению их значений на каждом шаге. Поэтому задачу оптимизации (10) на каждом шаге пришлось бы ставить заново с уточненными константами. В результате задача (10), рассчитанная на число шагов, необходимых для достижения конечной точности ε , превращалась бы каждый раз в одношаговую задачу, что снизило бы эффективность ее практического применения.

Рассмотрим функцию трудоемкости $C(\eta_k,\delta_k)$. Предположим, что эта функция является неубывающей по переменной η_k и невозрастающей по δ_k . Исходя из этого, а также из возрастания функции $\psi(\eta_k,\delta_k)$ по обеим переменным, нетрудно показать, что в целевой функции задачи (1) максимум по $\{\eta_k\}_{k=0}^{m-1}$ достигается при обращении неравенства (9) в равенство, т.е. при выполнении следующего условия:

$$\left.\begin{aligned} &\hat{\eta}_0 = \eta_0 , \\ &\hat{\eta}_{k+1} = \psi(\hat{\eta}_k, \gamma_k), k = 0,1,\ldots, \end{aligned}\right\} \qquad (11)$$

а минимум по $\{\gamma_k\}_{k=0}^{m-1}$ – при справедливости тождества $\eta_m = \varepsilon$.
Поэтому задача (10) примет вид

$$\left.\begin{aligned} &\sum_{k=0}^{m-1} C(\eta_k, \gamma_k) \longrightarrow min, \\ &\eta_{k+1} = min\{\rho; L_k \eta_k\}\eta_k + M_k \gamma_k, \\ &\eta_m = \varepsilon. \end{aligned}\right\} \qquad (12)$$

Далее докажем, что при наложении на функцию трудоемкости дополнительного естественного требования $C(\eta, \gamma) \geq \delta > 0$ последовательность $\{\eta_k\}_{k=0}^{m}$, соответствующая оптимальному выбору $\{\gamma_k\}_{k=0}^{m-1}$, может выбираться только среди убывающих последовательностей, т.е. таких, для которых справедливо $\eta_{k+1} < \eta_k$, $k = 0,\ldots,m-1$. Иными словами, нужно показать, что, имея η_0 – данное начальное приближение, последовательность $\{\gamma_k\}_{k=0}^{m-1}$ и соответствующую им по формуле $\eta_{k+1} = \psi(\eta_k, \gamma_k)$ неубывающую последовательность $\{\eta_k\}_{k=0}^{m}$, всегда можно построить новую последовательность $\{\gamma_k^*\}_{k=0}^{m-1}$ такую, что соответствующая ей последовательность $\{\eta_k^*\}_{k=0}^{m}$ будет убывать и будет выполняться неравенство

$$\sum_{k=0}^{m-1} C(\eta_k^*, \gamma_k^*) < \sum_{k=0}^{m-1} C(\eta_k, \gamma_k).$$

Предположим, первоначально имели последовательность $\{\eta_k\}_{k=0}^{m}$ такую, что

$$\eta_0 > \eta_1 > \ldots > \eta_{i-1} \leq \eta_i > \eta_{i+1} > \ldots > \eta_{j-1} \leq \eta_j > \eta_{j+1} > \ldots > \eta_m = \varepsilon .$$

Для построения новой последовательности $\{\eta_k^*\}$:

1. Найдем номер $\bar{i}$ такой, что $\bar{i} \geq i-1$ и $\bar{i}$ – ближайший к $i-1$, при котором $\eta_{i-1} > \eta_{\bar{i}}$. Если $\eta_{i-1} \leq \eta_i$, то $\bar{i} \geq i+1$. Проделаем это для всех индексов j таких, что $\eta_{j-1} \leq \eta_j$.

2. Для всех тех индексов j, для которых $\eta_{j-1} \leq \eta_j$, вычеркнем из начальной последовательности $\{\eta_k\}_{k=0}^{m}$ члены $\eta_j,\ldots,\eta_{\bar{j}-1}$. Поскольку $\bar{j}-1 \geq j$, то для каждого такого j множество вычеркнутых индексов непусто.

Для построения новой последовательности $\{\gamma_k^*\}_{k=0}^{m-1}$:

1. Для каждого вычеркнутого η_j вычерчивает соответственно

из начальной последовательности $\{\gamma_k\}$ член γ_{j-1} и из последовательности констант $\{L_k\}, \{M_k\}$ члены L_{j-1}, M_{j-1}.

2. Нумеруем полученные последовательности $\{\eta_k^*\}, \{\gamma_k^*\}$, $\{L_k^*\}, \{M_k^*\}$, начиная с 0.

3. Для каждого индекса j пересчитываем γ_j^* так, чтобы выполнялось неравенство $\eta_{j+1}^* = \min\{\rho, L_j^* \eta_j^*\}\eta_j^* + M_j^* \gamma_j^*$.

Теперь покажем, что

$$\sum_{k=0}^{m^*-1} C(\eta_k^*, \gamma_k^*) < \sum_{k=0}^{m-1} C(\eta_k, \gamma_k), \tag{13}$$

где m^* – наименьший номер, при котором $\eta_{m^*}^* = \varepsilon$. Очевидно, $m^* < m$. Распишем каждую из сумм:

$$\sum_{k=0}^{m-1} C(\eta_k, \gamma_k) = C(\eta_0, \gamma_0) + \ldots + C(\eta_{i-1}, \gamma_{i-1}) + C(\eta_i, \gamma_i) + C(\eta_{i+1}, \gamma_{i+1}) + \ldots$$

$$\sum_{k=0}^{m^*-1} C(\eta_k^*, \gamma_k^*) = C(\eta_0, \gamma_0) + \ldots + C(\eta_{i-1}, \gamma_{\bar{i}-1}^*) + \ldots,$$

где $\gamma_{\bar{i}-1}^*$ удовлетворяет неравенству

$$\eta_{\bar{i}} = \min\{\rho; L_{\bar{i}-1}\eta_{i-1}\}\eta_{i-1} + M_{\bar{i}-1}\gamma_{\bar{i}-1}^*. \tag{14}$$

Вместе с тем выполняется следующее неравенство, связывающее первоначальные величины $\eta_{\bar{i}}, \eta_{\bar{i}-1}, \gamma_{\bar{i}-1}$:

$$\eta_{\bar{i}} = \min\{\rho; L_{\bar{i}-1}\eta_{\bar{i}-1}\}\eta_{\bar{i}-1} + M_{\bar{i}-1}\gamma_{\bar{i}-1}. \tag{15}$$

В силу выборе индекса $\bar{i}$ имеем $\eta_{i-1} \le \eta_{\bar{i}-1}$, поэтому из равенств (14), (15) следует, что $\gamma_{\bar{i}-1}^* \ge \gamma_{\bar{i}-1}$. Так как из последнего неравенства получаем $C(\eta, \gamma_{\bar{i}-1}^*) \le C(\eta, \gamma_{\bar{i}-1})$, то для доказательства (13) достаточно доказать

$$C(\eta_{i-1}, \gamma_{\bar{i}-1}) + \ldots < C(\eta_{i-1}, \gamma_{i-1}) + C(\eta_i, \gamma_i) + C(\eta_{i+1}, \gamma_{i+1}) + \ldots$$

Далее, в силу выбора индекса $\bar{i}$ выполняется неравенство $\eta_{i-1} \le \eta_{\bar{i}-1}$, откуда следует справедливость соотношения $C(\eta_{i-1}, \gamma) \le C(\eta_{\bar{i}-1}, \gamma)$. Поэтому достаточно доказать справедливость неравенства

$$C(\eta_{\bar{i}-1}, \gamma_{\bar{i}-1}) + \ldots < C(\eta_{i-1}, \gamma_{i-1}) + C(\eta_i, \gamma_i) + C(\eta_{i+1}, \gamma_{i+1}) + \ldots \tag{16}$$

Так как $\bar{i} - 1 \ge i$, то $\bar{i} - 1$ – один из элементов множества индексов $\{i, i+1, \ldots\}$, т.е. в правой части неравенства (16) содержатся все слагаемые из его левой части. Кроме того, в правой части обязательно есть слагаемые, которых нет слева.

Например, таким является слагаемое $C(\eta_{i-1}, \gamma_{i-1})$. Поскольку $C(\eta_{i-1}, \gamma_{i-1}) > 0$, то неравенство (16) справедливо, следовательно, выполняется (13), что означает убывание последовательности $\{\eta_k\}$, соответствующей оптимальной $\{\gamma_k^*\}$.

Будем предполагать, что наш внутренний процесс сходится с линейной скоростью, т.е. существует $w < 1$ такое, что

$$\|z^t - y^{k+1}\| \leq w^t \|z^0 - y^{k+1}\|, \text{ где } z^0 = x^k. \tag{17}$$

Поэтому число итераций, которое достаточно сделать для получения из точки z^0 точки z^t, равно t. Допустим, t – наименьший номер, при котором справедливо неравенство $\|h_k(z^t)\| \leq \gamma_k$. Очевидно, $t > 0$. Найдем функцию трудоемкости на k-м шаге как функцию, зависящую от параметров η_k, γ_k и являющуюся верхней границей для t. Применяя оценку (7) для функции $\|h_k(x)\|$ к неравенству (17), получим

$$t - 1 \leq \log_w \frac{\|z^{t-1} - y^{k+1}\|}{\|z^0 - y^{k+1}\|} \leq \log_w \frac{1/b^k \|h_k(z^{t-1})\|}{1/a^k \|h_k(z^0)\|} <$$

$$< \log_w \frac{a^k \gamma_k}{b^k \|h_k(x^k)\|} = \log_w \left(\frac{a^k}{b^k} \frac{\gamma_k}{\eta^k} \right).$$

Отсюда

$$t < \log_w \frac{\gamma_k}{\eta_k} + \log_w \frac{a^k}{b^k} + 1 = \frac{1}{\ln(1/w)} \left(\ln \frac{\eta_k}{\gamma_k} + \ln \frac{b^k}{a^k w} \right), \tag{18}$$

где a^k, b^k – константы из оценки (7) для функции $\|h_k(x)\|$, которые, очевидно, меняются на каждом внешнем шаге k. Для упрощения оценим их через a, b – соответствующие константы для функции $\|h(x)\|$. Нетрудно проверить справедливость соотношений $\begin{cases} a \leq a^k \\ b \geq b^k \end{cases}$, подставляя которые в неравенство (18), получим

$$0 \leq t < \frac{1}{\ln(1/w)} \left(\ln \frac{\eta_k}{\gamma_k} + \ln \frac{b}{aw} \right),$$

откуда заключаем

$$C(\eta_k, \gamma_k^*) = \frac{1}{\ln(1/w)} \left(\ln \frac{\eta_k}{\gamma_k} + \ln \frac{b}{ab} \right) > 0. \tag{19}$$

Таким образом, задача (12) принимает вид

$$\begin{cases} \sum\limits_{k=0}^{m-1} \ln\left(\frac{\eta_k}{\gamma_k}\right) + m\cdot\sigma \longrightarrow \min\limits_{\gamma_k>0,\, m>0}, \text{ где } \sigma = \ln\frac{b}{aw} > 0; \\ \eta_{k+1} = \min\{\rho; \Delta_k\eta_k\}\eta_k + M_k\gamma_k,\ k=0,1,\ldots,m-1; \\ \eta_m = \varepsilon. \end{cases} \quad (20)$$

Заметим, что функция трудоемкости (I9) обладает всеми указанными выше свойствами, именно, положительностью, возрастанием по аргументу η_k и убыванием по γ_k, поэтому при оптимальном выборе последовательности $\{\gamma_k\}_{k=0}^{m-1}$ выполняется $\eta_{k+1} < \eta_k,\ k=0,\ldots,m-1$. Заметим также, что последовательность $\{\Delta_k\}_{k=0}^{m-1}$ является невозрастающей, значит, $\{\Delta_k\eta_k\}_{k=0}^{m-1}$ будет убывать. Поэтому, как в работе [6], для решения задачи (20) достаточно рассмотреть три случая:

1) $\rho \geq \Delta_0\eta_0 > \varepsilon$;
2) $\Delta_0\eta_0 > \varepsilon \geq \rho$;
3) $\Delta_0\eta_0 > \rho > \varepsilon$.

Причем более общим является случай 3), в котором аналогично соответствующему случаю, рассмотренному в работе [6], оптимальная трудоемкость смешанного процесса равна $S_1(\eta_0, \rho/\Delta_\ell) + S_2(\rho/\Delta_\ell, \varepsilon)$, где $S_i(\eta_1, \eta_2)$ – оптимальная трудоемкость процесса соответственно при линейной и квадратичной скорости убывания функции ошибки η_k с ее начальным значением η_1 и конечным η_2, ℓ – наименьший номер, при котором $\Delta_\ell\eta_\ell \leq \rho$.

Сначала рассмотрим подзадачу, соответствующую квадратичному убыванию функции η_k:

$$\left.\begin{aligned} &\sum_{k=\ell}^{m-1} \ln\left(\frac{\eta_k}{\gamma_k}\right) + (m-\ell)\sigma \longrightarrow \min_{\gamma_k>0,\, m>\ell}; \\ &\eta_{k+1} = \Delta_k\eta_k^2 + M_k\gamma_k,\quad k=\ell,\ldots,m-1; \\ &\eta_m = \varepsilon. \end{aligned}\right\} \quad (21)$$

Подставив управляющий параметр γ_k в виде

$$\gamma_k = \frac{(1+\xi_k)\Delta_k\eta_k^2}{M_k}, \quad (22)$$

получим равенство, связывающее конечную погрешность η_{k+1} k-го шага с его начальной погрешностью η_k:

$$\eta_{k+1} = (1+\xi_k)\angle_k \eta_k^2. \tag{23}$$

Подставив соотношения (22), (23) в (21), получим следующую задачу:

$$\left.\begin{aligned} &\ln\left(\frac{\eta_\ell}{\varepsilon}\prod_{k=\ell}^{m-1} M_k \frac{1+\xi_k}{\xi_k}\right) + (m-\ell)\sigma \to \min_{\xi_k>0,\, m>\ell};\\ &\eta_{k+1} = (1+\xi_k)\eta_k^2\angle_k,\ k=\ell,\dots,m-1;\\ &\eta_m = \varepsilon. \end{aligned}\right\} \tag{24}$$

Аналогично тому, как это делается в работе [6], можно показать, что минимизируемая функция на своем множестве ограничений принимает наименьшее значение при $\xi_k = 2^{k-\ell}\xi_\ell,\ k=\ell,\dots,m-1$. С учетом этих равенств задача (24) переписывается следующим образом:

$$\left.\begin{aligned} &\ln\frac{\eta_\ell}{\varepsilon}\prod_{k=\ell}^{m-1}\frac{(1+2^{k-\ell}\xi_\ell)M_k}{2^{k-\ell}\xi_\ell} + (m-\ell)\sigma \to \min_{\xi_\ell>0,\, m>0};\\ &\eta_{k+1} = (1+2^{k-\ell}\xi_\ell)\eta_k^2\angle_k,\ k=\ell,\dots,m-1;\\ &\eta_m = \varepsilon. \end{aligned}\right\} \tag{25}$$

С целью придать задаче (25) вид, более удобный для решения, вводятся бесконечные произведения

$$P(\xi_k) = \prod_{i=0}^{\infty}\left[(1+2^i\xi_k)\angle_{k+i}\right]^{\frac{1}{2^{i+1}\xi_k}};\quad \widetilde{P}(\xi_k) = \prod_{i=0}^{\infty}\frac{1+2^i\xi_k}{2^i\xi_k} M_{k+i}.$$

Тогда (25) формулируется следующим образом:

$$\left.\begin{aligned} &\ln\left(\frac{\eta_\ell}{\varepsilon}\cdot\frac{\widetilde{P}(\xi_\ell)}{\widetilde{P}(2^{m-\ell}\xi_\ell)}\right) + (m\ \ell)\sigma \to \min_{\xi_\ell>0,\, m>\ell};\\ &\left[\eta_\ell^{\frac{1}{\xi_\ell}}\frac{P(\xi_\ell)}{P(2^{m-\ell}\xi_\ell)}\right]^{2^{m-\ell}\xi_\ell} = \varepsilon. \end{aligned}\right\} \tag{26}$$

Прологарифмировав (26), получим систему

$$\left.\begin{aligned} &\ln\widetilde{P}(\xi_\ell) - \ln\widetilde{P}(2^{m-\ell}\xi_\ell) + (m-\ell)\sigma \to \min_{\xi_\ell>0,\, m>\ell};\\ &2^{m-\ell}\ln\eta_\ell + 2^{m-\ell}\xi_\ell\ln P(\xi_\ell) - 2^{m-\ell}\xi_\ell\ln P(2^{m-\ell}\xi_\ell) = \ln\varepsilon. \end{aligned}\right\} \tag{27}$$

Чтобы решить задачу (27) стандартным методом, находящим точку экстремума гладкой функции с гладкими ограничениями, обобщим задачу (27) на случай произвольного вещественного m. Так как произвольную точку $y \in R^1, y>0$ можем представить в виде $y = \xi_z = 2^{z-\ell}\xi_\ell$, то обобщенные произведения будут выглядеть следующим образом:

$$P(y) = \prod_{i=0}^{\infty}\left[(1+2^i y)L_{z+i}\right]^{\frac{1}{2^{i+1}y}};\quad \widetilde{P}(y) = \prod_{i=0}^{\infty}\left(\frac{1+2^i y}{2^i y}M_{z+i}\right),$$

где L_t, M_t, $t \in R^n$, – функции, обобщающие последовательности $\{L_k\}$ и $\{M_k\}$ так, что график каждой из этих функций представляет собой гладкую кривую, проходящую соответственно через точки L_k, M_k, $k = \ell, \ldots, m-1$. Так как функции $L_t, M_t, t \in R^n$, считаем заданными заранее, то будем предполагать их независимость от переменной y. Тогда справедливо следующее утверждение:

$$y(\ln\widetilde{P}(y))' = -(2y\ln P(y))', \tag{28}$$

которое ввиду ограниченности функций L_t, M_t доказывается аналогично лемме 3 [6]. Далее, если в (28) переменную y заменить на $\sigma = 2^{m-\ell}y$, то при взятии производной по y или по m равенство (28) сохранится, так как множители дифференцирования σ'_y или σ'_m взаимно сократятся. Таким образом, справедливы следующие равенства:

$$2^{m-\ell}y(\ln\widetilde{P}(2^{m-\ell}y))'_y = -(2^{m-\ell+1}y\ln P(2^{m-\ell}y))'_y; \tag{29}$$

$$2^{m-\ell}y(\ln\widetilde{P}(2^{m-\ell}y))'_m = -(2^{m-\ell+1}y\ln P(2^{m-\ell}y))'_m. \tag{30}$$

С учетом сказанного применим правило Лагранжа к задаче (27):

$$\left.\begin{aligned}
&(\ln\widetilde{P}(\xi_\ell) - \ln\widetilde{P}(2^{m-\ell}\xi_\ell))'_{\xi_\ell} + \\
&+\lambda(2^{m-\ell}\xi_\ell\ln P(\xi_\ell) - 2^{m-\ell}\xi_\ell\ln P(2^{m-\ell}\xi_\ell))'_{\xi_\ell} = 0; \\
&(-\ln\widetilde{P}(2^{m-\ell}\xi_\ell) + (m-\ell)\sigma)'_m + \\
&+\lambda(2^{m-\ell}\ln\eta_\ell + 2^{m-\ell}\xi_\ell\ln P(\xi_\ell) - 2^{m-\ell}\xi_\ell\ln P(2^{m-\ell}\xi_\ell))'_m = 0.
\end{aligned}\right\}$$

Применив (29) и (30) к первому уравнению системы, получим

$$(\ln\tilde{P}(\xi_e))'-(\ln\tilde{P}(2^{m-e}\xi_e))'-\lambda 2^{m-e-1}\xi_e(\ln\tilde{P}(\xi_e))'+$$
$$+\lambda 2^{m-e-1}\xi_e(\ln\tilde{P}(2^{m-e}\xi_e))'=0.$$

Отсюда $(1-\lambda 2^{m-e-1}\xi_e)\cdot[(\ln\tilde{P}(\xi_e))'-(\ln\tilde{P}(2^{m-e}\xi_e))']=0$,
т.е. $(1-\lambda 2^{m-e-1}\xi_e)=0$ или $[(\ln\tilde{P}(\xi_e))'-(\ln\tilde{P}(2^{m-e}\xi_e))']=0$.
Покажем, что выражение в квадратной скобке не может быть равным 0 при $m>e$. Действительно,

$$(\ln\tilde{P}(\xi_e))'=-\sum_{i=0}^{\infty}\frac{1}{(1+2^{i}\xi_e)\xi_e},$$

$$(\ln\tilde{P}(2^{m-e}\xi_e))'=-\sum_{i=0}^{\infty}\frac{1}{(1+2^{i+m-e}\xi_e)\xi_e}.$$

поэтому $(\ln\tilde{P}(\xi_e))'-(\ln\tilde{P}(2^{m-e}\xi_e))'=0$ тогда и только тогда, когда $m-e=0$. Но считаем $m>e$, отсюда $(1-\lambda 2^{m-e-1}\xi_e)=0$, следовательно,

$$\lambda=\frac{1}{2^{m-e-1}\xi_e}.$$

Теперь применим (30) ко второму уравнению системы:

$$(-\ln\tilde{P}(2^{m-e}\xi_e))'+\sigma+\lambda 2^{m-e}\ln 2\ln\eta_e+\lambda 2^{m-e}\ln 2\,\xi_e P(\xi_e)+$$
$$+\lambda 2^{m-e-1}\xi_e(\ln\tilde{P}(2^{m-e}\xi_e))'=0.$$

Отсюда

$$[\ln\tilde{P}(2^{m-e}\xi_e)]'\times(-1+\lambda 2^{m-e-1}\xi_e)+$$
$$+\lambda 2^{m-e}\ln 2(\ln\eta_e+\xi_e P(\xi_e))+\sigma=0.$$

Подставляя значение λ, получим $\frac{2\ln 2}{\xi_e}(\ln\eta_e+\xi_e P(\xi_e))+\sigma=0.$
Домножая последнее равенство на $\frac{\xi_e}{2\ln 2}$ и вычитая из второго уравнения (27) это равенство, умноженное на $\frac{2^{m-e-1}\xi_e}{\ln 2}$, получим следующую систему уравнений:

$$\left.\begin{aligned}&\ln\eta_e+\xi_e\ln P(\xi_e)+\frac{\xi_e\sigma}{2\ln 2}=0,\\&\ln\varepsilon+2^{m-e}\xi_e\ln P(2^{m-e}\xi_e)+\frac{2^{m-e-1}\xi_e\sigma}{\ln 2}=0.\end{aligned}\right\}\qquad(31)$$

Для исследования системы (3I) на существование решения распишем ее более подробно:

$$\left.\begin{aligned}
&\ln\eta_\ell+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln(1+2^{i}\xi_\ell)+\\
&+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln\Delta_{i+\ell}+\frac{\xi_\ell\sigma}{2\ln 2}=0;\\
&\ln\varepsilon+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln(1+2^{i+m-\ell}\xi_\ell)+\\
&+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln\Delta_{i+m-\ell}+\frac{2^{m-\ell-1}\xi_\ell\sigma}{\ln 2}=0.
\end{aligned}\right\}\qquad(32)$$

Функция, стоящая в левой части первого уравнения системы, является возрастающей по переменной ξ_ℓ , стремится к $+\infty$ при $\xi_\ell\to+\infty$, а при $\xi_\ell=0$ она принимает значение $\ln\eta_\ell+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln\Delta_{i+\ell}$. В силу невозрастания последовательности $\{\Delta_k\}$

$$\ln\eta_\ell+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln\Delta_{i+\ell}\le\ln\eta_\ell+$$
$$+\ln\Delta_\ell\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}=\ln(\eta_\ell\Delta_\ell)<0,$$

что означает существование решения $\xi_\ell^*>0$ и его единственность.

Рассмотрим теперь второе уравнение системы (32). Как уже указывалось, предполагаем ограниченность функции Δ_t снизу, т.е. для всех $t\in R$, $\Delta_t\ge\Delta_{min}>0$. Тогда, обозначив правую часть рассматриваемого уравнения за $\psi(m)$, при $m\to\infty$ имеем

$$\psi(m)\ge\ln\varepsilon+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln(1+2^{i+m-\ell}\xi_\ell)+$$
$$+\ln\Delta_{min}+\frac{2^{m-\ell-1}\xi_\ell\sigma}{\ln 2}\longrightarrow\infty.$$

При $m=\ell$

$$\psi(l)=\ln\varepsilon+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln(1+2^i\xi_l)+$$

$$+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln\Delta_{i+l}+\frac{\xi_l\sigma}{2\ln 2}<\ln\eta_l+$$

$$+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln(1+2^i\xi_l)+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln\Delta_{i+l}+\frac{\xi_l\sigma}{2\ln 2}=0,$$

так как последнее равенство есть в точности первое уравнение (32). То есть $\psi(l)<0$, откуда на основании непрерывности $\psi(m)$ следует существование $m^*>l$ такого, что $\psi(m^*)=0$.

Таким образом, для рассмотренного случая задача оптимального управления разрешима, причем для нахождения параметра ξ_l, а следовательно, и для вычисления управляющего параметра $\gamma_l=\frac{\Delta_l\xi_l\eta_l^2}{M_l}$ необходимо знать лишь значение погрешности текущего шага η_l, в то время как априорное знание конечной точности ε позволяет определить число шагов m, необходимых для нахождения решения нелинейной задачи дополнительности с заданной точностью.

Теперь рассмотрим подзадачу, соответствующую линейной скорости убывания функции ошибки η_k. Задача (20) принимает вид

$$\left.\begin{aligned}&\sum_{k=0}^{l-1}\ln\left(\frac{\eta_k}{\gamma_k}\right)+l\sigma\longrightarrow\min_{\gamma_k>0,\,l>0},\\&\eta_{k+1}=\rho\eta_k+M_k\gamma_k,\ k=0,1,\ldots,l-1,\\&\eta_l=\delta,\end{aligned}\right\}\qquad(33)$$

где δ выбирается из условия $\eta_l\Delta_l=\rho$, т.е. $\delta=\frac{\rho}{\Delta_l}$. Представив γ_k в виде $\gamma_k=\frac{\rho\xi_k\eta_k}{M_k}$, получим $\eta_{k+1}=(1+\xi_k)\rho\eta_k$, и задача (33) перепишется таким образом:

$$\left.\begin{aligned}&\ln\left(\frac{\eta_0}{\delta}\prod_{k=0}^{l-1}M_k\frac{1+\xi_k}{\xi_k}\right)+l\sigma\longrightarrow\min_{\xi_k>0,\,l>0},\\&\eta_{k+1}=(1+\xi_k)\rho\eta_k,\ k=0,\ldots,l-1,\\&\eta_l=\delta.\end{aligned}\right\}\qquad(34)$$

Аналогично квадратичному случаю найдем связь между величинами ξ_k, $k=0,1,\ldots,\ell-1$, получив при этом $\xi_k=\xi_0$ для всех $k=0,1,\ldots,\ell-1$. Задача (34), таким образом, преобразуется в систему

$$\left.\begin{aligned} &\ln\left[\frac{\rho_0}{\delta}\prod_{k=0}^{\ell-1}M_k\left(\frac{1+\xi_0}{\xi_0}\right)^{\ell}\right]+\ell\sigma\longrightarrow\min_{\xi_0>0,\,\ell>0}, \\ &\rho_{k+1}=(1+\xi_0)\rho\rho_0,\quad k=0,\ldots,\ell-1, \\ &\rho_\ell=\delta, \end{aligned}\right\} \qquad (35)$$

которая решается аналогично задаче (26), именно, последовательность $\{M_k\}$ обобщается гладкой функцией M_t, $t\in R$, после чего система (35) логарифмируется и решается по правилу Лагранжа. В результате получаем задачу

$$\left.\begin{aligned} &\ln M_{\ell-1}+\ln(1+\xi_0)-\ln\xi_0+\frac{1}{\xi_0}\ln(1+\xi_0)+\frac{1}{\xi_0}\ln\rho+\sigma=0, \\ &\ell\ln(1+\xi_0)+\ell\ln\rho+\ln\rho_0=\ln\delta. \end{aligned}\right\} \qquad (36)$$

Величина ξ_0 находится из первого уравнения системы, домножив которое на ξ_0, получим

$$F(\xi_0)=\xi_0(\ln M_{\ell-1}+\sigma)+\xi_0\ln\frac{1+\xi_0}{\xi_0}+\ln(1+\xi_0)+\ln\rho=0. \qquad (37)$$

Исследуем уравнение (37) на существование решения. $F'(\xi_0)=\ln M_{\ell-1}+\sigma+\ln\left(1+\frac{1}{\xi_0}\right)>0$ при $\xi_0>0$, что означает монотонное возрастание функции $F(\xi_0)$ при $\xi_0>0$. Далее, пользуясь правилом Лопиталя, получим

$$\lim_{\xi_0\to 0}F(\xi_0)=\ln\rho<0;\quad \lim_{\xi_0\to\infty}F(\xi_0)=\infty.$$

Следовательно, в силу непрерывности $F(\xi_0)$ существует единственное ξ_0^* такое, что $F(\xi_0^*)=0$.

Из второго уравнения системы (36) найдем $\ell>0$ – число шагов, которое необходимо сделать, чтоб выполнилось условие $\angle_\ell\rho_\ell=\rho$. Получаем $\ell=\dfrac{\ln\delta-\ln\rho_0}{\ln(1+\xi_0^*)+\ln\rho}$. Дробь, стоящая в правой части последнего равенства, имеет отрицательный числитель. Поэтому для существования $\ell>0$ необходимо, чтоб выполнилось неравенство $\ln(1+\xi_0^*)+\ln\rho<0$, что равносильно $\xi_0^*<\dfrac{1-\rho}{\rho}$. Заметим, что

$$F\left(\frac{1-\rho}{\rho}\right)=\frac{1-\rho}{\rho}\ln M_{\ell-1}+\frac{1-\rho}{\rho}\ln\frac{1}{1-\rho}>0,$$

так как $\rho<1$. Исходя из монотонного возрастания функции $F(\xi_0)$ и выполнения равенства $F(\xi_0^*)=0$, получаем $\xi_0^*<\frac{1-\rho}{\rho}$.

Таким образом, второе уравнение системы (36) также однозначно разрешимо.

2. Практическая реализация управления неточным методом Ньютона

Рассмотрим теперь конкретную реализацию неточного метода Ньютона для решения задачи (I) с применением описанной выше схемы.

Пусть имеется начальное приближение $x^0\in R^n$ и приближения L_0, M_0 констант L и M. В более общем случае $L_0\eta_0>\rho$. Тогда точность γ_0 определяется из равенства $\gamma_0=\frac{\rho\xi_0\eta_0}{M_0}$, где ξ_0 находится как решение уравнения (37). Затем отыскивается точка $x^1\in R^n$ такая, что $\|h_0(x^1)\|\le\gamma_0$, и определяется новая погрешность $\eta_1=\|h(x^1)\|$. Завершается шаг пересчетом константы M_0, именно, из равенства $\eta_1=\rho\eta_0+\tilde M_0\gamma_0$ определяется константа $\tilde M_0$, после чего для заданного $\alpha\in(0;1)$ полагается $M_1=\alpha M_0+(1-\alpha)\max\{0;\tilde M_0\}$. Процесс этот продолжается до тех пор, пока на некотором шаге ℓ не выполнится неравенство $L_0\eta_\ell\le\rho$. Далее переходим к квадратичной подзадаче, полагая при этом $L_\ell=L_0$. Уравнение для определения ξ_ℓ в этом случае имеет вид

$$\ln\eta_\ell+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln(1+2^i\xi_\ell)+\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln L_{i+\ell}+\frac{\xi_\ell G}{2\ln 2}=0. \qquad (38)$$

Введем обозначения $\ln\bar L_\ell=\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln L_{i+\ell}$. Так как при практическом решении задачи (I) константы L_i, $i=\ell,\ell+1,\dots$, заранее не известны, то точно определить $\bar L_\ell$ не можем. Однако имеем $L_{i+1}\le L_i, i=\ell,\ell+1,\dots$, поэтому $\bar L_\ell\le L_\ell$. Выбрав $\bar L_\ell$ из последнего неравенства и подставив его в уравнение (38), найдем параметр ξ_ℓ как решение этого уравнения, после чего определим точность γ_ℓ из равенства

$\gamma_\ell = \dfrac{L_\ell \xi_\ell \eta_\ell^2}{M_\ell}$, рассматривая при этом γ_ℓ как функцию от L_ℓ : $\gamma_\ell = \gamma_\ell(L_\ell)$. Затем, найдя точку $x^{\ell+1} \in R^n$ такую, что $\|h_\ell(x^{\ell+1})\| \le \gamma_\ell$, и определив функцию ошибки $\eta_{\ell+1} = \|h(x^{\ell+1})\|$, вычислим константы $L_{\ell+1}, M_{\ell+1}, \bar{L}_{\ell+1}$. Имея точку $x^{\ell+1}$ и погрешность $\eta_{\ell+1}$, нахождение $L_{\ell+1}$ производится таким образом, чтобы, не нарушая условия $\|h_\ell(x^{\ell+1})\| \le \gamma_\ell$, неравенство, связывающее $\eta_{\ell+1}$ с величиной $(1+\xi_\ell)\eta_\ell^2 L_\ell$, максимально приблизить к равенству, соблюдая при этом оценку $L_{\ell+1} \le L_\ell$.

Для вычисления $L_{\ell+1}$ найдем $\tilde{\tilde{L}}_\ell$ из равенства $\|h_\ell(X^{\ell+1})\| = \gamma_\ell(\tilde{\tilde{L}}_\ell)$, иными словами, $\tilde{\tilde{L}}_\ell$ – предельная величина, при которой неравенство обращается в равенство, т.е.

$$\tilde{\tilde{L}}_\ell = \min\{\hat{L} : \|h_\ell(X^{\ell+1})\| \le \gamma_\ell(\hat{L})\}, \tag{39}$$

откуда $\tilde{\tilde{L}}_\ell \le L_\ell$ и $\tilde{\tilde{L}}_\ell = \dfrac{\|h_\ell(X^{\ell+1})\| M_\ell}{\xi_\ell \eta_\ell^2}$.

Затем находим $\tilde{L}_\ell$ из равенства $\eta_{\ell+1} = (1+\xi_\ell)\tilde{L}_\ell \eta_\ell^2$, откуда следует, что

$$\tilde{L}_\ell = \frac{\eta_{\ell+1}}{(1+\xi_\ell)\eta_\ell^2} . \tag{40}$$

В данной ситуации возможны два случая.

1. $\tilde{L}_\ell \ge \tilde{\tilde{L}}_\ell$. Тогда, в соответствии с (39), выполняется неравенство $\|h_\ell(x^{\ell+1})\| \le \gamma_\ell(\tilde{L}_\ell)$, поэтому принимаем $L_{\ell+1} = \min\{L_\ell ; \tilde{L}_\ell\}$.

2. $\tilde{L}_\ell < \tilde{\tilde{L}}_\ell$. Тогда имеем $\tilde{L}_\ell < \tilde{\tilde{L}}_\ell \le L_\ell$, следовательно, $\tilde{\tilde{L}}_\ell$ – ближайшее к $\tilde{L}_\ell$ значение, при котором выполняется условие $\|h_\ell(x^{\ell+1})\| \le \gamma_\ell(\tilde{\tilde{L}}_\ell)$. На основании этого полагаем $L_{\ell+1} = \tilde{\tilde{L}}_\ell$. Для вычисления константы $M_{\ell+1}$ сначала найдем $\tilde{M}_\ell$ из равенства $\eta_{\ell+1} = \tilde{L}_\ell \eta_\ell^2 + \tilde{M}_\ell \gamma_\ell$, используя (40), получив при этом $\tilde{M}_\ell = \dfrac{\eta_{\ell+1} \xi_\ell}{\gamma_\ell(1+\xi_\ell)}$. Затем, задав $\alpha \in (0;1)$, положим $M_{\ell+1} = \alpha M_\ell + (1-\alpha)\tilde{M}_\ell$.

Константу $\bar{L}_{\ell+1}$ найдем из следующих соображений. При нахождении ξ_ℓ в уравнении (38) имели слагаемое

$$\ln \bar{L}_\ell = \sum_{i=0}^{\infty} \frac{1}{2^{i+1}} \ln L_{\ell+i} = \frac{1}{2}\ln L_\ell + \frac{1}{4}\ln L_{\ell+1} + \frac{1}{8}\ln L_{\ell+2} + \ldots \tag{41}$$

Параметр ξ_{i+1} находится из аналогичного уравнения, отличающегося от (38) лишь заменой индекса ℓ на $\ell+1$. Поэтому слагаемое, содержащее константы L_i, $i=\ell+1,\ell+2,\dots$, будет иметь вид

$$\ln \bar{L}_{\ell+1}=\sum_{i=0}^{\infty}\frac{1}{2^{i+1}}\ln L_{\ell+1+i}=\frac{1}{2}\ln L_{\ell+1}+\frac{1}{4}\ln L_{\ell+2}+\frac{1}{8}\ln L_{\ell+3}+\dots \qquad (42)$$

Из равенств (41), (42) получаем $\ln\bar{L}_{\ell+1}=2\ln\bar{L}_{\ell}-\ln L_{\ell}$. Следовательно, $\bar{L}_{\ell+1}=\frac{\bar{L}_\ell^2}{L_\ell}$. Но ввиду того, что L_ℓ уточнили константой $L_{\ell+1}$, полагаем $\bar{L}_{\ell+1}=\frac{\bar{L}_\ell^2}{L_{\ell+1}}$. С найденными ранее $x^{\ell+1}$, $L_{\ell+1}$, $M_{\ell+1}$, $\bar{L}_{\ell+1}$ процедура, описанная выше, повторяется. Если, наконец, для некоторого $m\in N$ окажется, что $\|h(x^m)\|\le\varepsilon$, процесс завершается, а точка x^m берется за приближенное решение x^* задачи (I).

В заключение укажем, каким способом можно отыскивать точку x^{k+1}, удовлетворяющую условию $\|h_k(x^{k+1})\|\le\gamma_k$. При минимизации самой функции $\|h_k(x)\|$ на R^n возникают трудности, связанные с тем, что эта функция лишь кусочно-гладкая. В силу этого предлагается минимизировать на R^n гладкую функцию

$$g(x)=\sum_{j=1}^{n}\left[(x_j)_-^2+(\omega_j)_-^2+(x_j)_+^2(\omega_j)_+^2\right],$$

где $\omega=f(x^k)+\nabla f(x^k)(x-x^k)$, $a_-=\min\{0;a\}$, $a_+=\max\{0;a\}$ для $a\in R^1$. Возможность отыскания точки x^{k+1} таким способом доказывается леммой в работе [5].

3. Результаты численных экспериментов

Численная реализация управления неточным методом Ньютона осуществлялась на трех тестовых функциях. Сравнение результатов проводилось с неточным методом Ньютона без управления. Критериями эффективности служили следующие величины:

I) N – число шагов, которое потребовалось для достижения заданной точности;

2) KO – общее число обращений к функции $g(x)$ за весь период работы;

3) $\alpha=\frac{KO}{\lg\eta_0-\lg\eta_m}$ – показатель, отражающий средние затраты вычислений $g(x)$ для понижения значения функции

ошибки $\rho_k = \|h(x^k)\|$ на один порядок. В качестве конечной точности для всех функций используется $\varepsilon = 10^{-2}$.

ПРИМЕР I. $f: R^2 \to R^2;\ f_1(x) = x_1 + x_2^2;\ f_2(x) = x_2$.
Начальное приближение $x^0 = (5;5)$.
Решение задачи дополнительности $x^* = (0;0)$.
Значение функции в искомой точке $f(x^*) = (0;0)$.

	N	KO	α
без управления	7	I2845	374I
с управлением	6	624	I9I

ПРИМЕР 2. $f: R^2 \to R^2;\ f_1(x) = x_1 - x_2;\ f_2(x) = x_1 x_2 + 1 - e^{-x_2}$.
Начальное приближение $x^0 = (5;5)$.
Решение задачи дополнительности $x^* = (0;0)$.
Значение функции в искомой точке $f(x^*) = (0;0)$.

	N	KO	α
без управления	I3	84I7	3064
с управлением	I2	3793	I372

ПРИМЕР 3. $f: R^5 \to R^5;\ f_i(x) = x_i + \sum_{j=i+1}^{5} x_j^2,\ i = 1,\ldots,5$.
Начальное приближение $x^0 = (5;5;5;5;5)$.
Решение задачи дополнительности $x^* = (0;0;0;0;0)$.
Значение функции в искомой точке $f(x^*) = (0;0;0;0;0)$.

	N	KO	α
без управления	43	3I3I3	976I
с управлением	33	I44II	4685

Литература

I. Megiddo N., Kojima M. On the existence and uniqueness of solution in nonlinear complementarity theory// Math. Program. - 1977.- V.12, N 1.- P.110-130.

2. Karamardian S. The nonlinear complementarity problem with Applications. Part I// J. Optim. Theory and Appl. - 1969.- V.4, N 2. - P. 87-98.

3. Pang J.S., Chan D. Iterative methods for variational and complementarity problems// Math. Program, - 1982. - V.24, N 3. - P. 284-311.

4. Pang J.S. Inexact Newton methods for nonlinear complementarity problem// Math. Program. - 1986. - V.36, N 1. - P.54-71.

5. Калашников В.В., Калашникова Н.И. Глобальная сходимость неточного метода Ньютона для решения нелинейной задачи дополнительности// Оптимизация. - 1988. - Вып. 42(59). - С.66-85.

6. Калашников В.В., Калашникова Н.И. Оптимальное управление внутренней точностью в двухуровневом итерационном процессе// Оптимизация. - 1988. - Вып.44(61). - С.27-55.

Поступила в ред.-изд. отдел
15.02.1989 г.